

文章编号:1005-3085(2004)07-0071-05

奥运场馆中临时商业网点设计中的 数学模型化方法

孟大志

(北京工业大学, 北京 100022)

摘 要: 本文简述2004年全国大学生数学建模竞赛A题的命题过程以及题目的立意与设计, 给出了该题命题人的解题思路与基本解法, 并对选择该题的大学生的论文进行介绍和评论。

1 引言

2001年中国申奥成功, 举国上下一片欢腾。当时我在香港中文大学访问, 在电视中看到这一消息时就想, 我能为2008年北京奥运会做些什么呢? A题的产生基于这样一种环境, 这样一种热情。2002年在大学生数学建模竞赛全国组委会的老师的支持下, 我开始构思这一课题。当时的目的, 不仅是为全国大学生数学建模比赛出一道贴近时事的赛题, 也希望通过全国大学生的集体智慧, 为北京奥运会“科技奥运”目标贡献一份力量, 那时候还只是希望, 通过本届参赛大学生的共同努力, 现在我们已经可以拿出有充分科学依据且切实可行的设计提供给北京奥组委。我想, 参赛的大学生们因此可以自豪地说: 我们也在为北京奥运会做贡献!

在筹备和举办奥运会的工作中, 科技奥运的份量应当是很重的。这不仅是表现在奥运场馆的建设, 媒介工作, 各个比赛项目的器材, 周边环境与整个城市建设, 以及开、闭幕式的表演等显而易见的“硬件”建设中, 而且也必然要表现在整个筹备与比赛实施全过程的管理工作中。这是一个复杂系统的管理和控制问题, 它不仅包括每项工作的细微的管理, 也包括整个系统的宏观调控, “管理出效益”将在这项长达六年的奥运筹备工作中充分体现其作用。而这个复杂系统的科学管理的核心是数字化的信息管理, 因此数学、信息科学与管理科学的结合将在其中发挥不可或缺的作用。但是, 这样一个信息管理并不是所有管理人员都能理解与付诸实施的, 本赛题设计的另外一个初衷就是想通过涉及面很大的大学生数学建模竞赛, 针对一个奥运会筹备过程中的具体的管理规划和设计的问题, 向奥组委提供“科学管理奥运”的一个建议。

出于以上初衷, 本赛题在2002年已经开始设计, 恰逢2003年的SARS, 为配合当年的时事, 把这个题目又搁置下来。在这期间我们又精化题目, 在北京市市政管委会和北京市城市规划委员会的支持下, 我们在题目中加入了真实的奥运主场馆规划图和“鸟巢”和“水立方”等独具特色的主场馆设计图。这项工作使远离北京的大学生们更能体会到该题与奥运会的密切联系, 激发同学们通过数模比赛为2008年北京奥运作一份贡献的热情。

2004年, 这道酝酿了许久的奥运竞赛题终于和参赛大学生们见面了, 近一半的大学生选择了这一道题, 通过评卷, 让我们高兴地感觉到大学生们对北京奥运会的积极支持与参与的热情, 一些论文向北京奥组委如何提高科技奥运的作用提出建议, 更多的论文向北京奥组委表达敬意与祝愿, 使论文充满了爱奥运, 爱祖国的感情。通过评卷, 我们看到中国大学生的聪明才智, 众多试卷表现出十分独特的观念, 十分聪明的技巧和比较丰富的数学建模的思想。一些方法的提出了超出了我们的预想, 使判卷老师十分惊叹: 后生可畏, 后生可敬! 总之, 该题的竞赛结果已经超出了预期地实现了命题者的初衷, 也为“北京奥运”筹备工作主动作出贡献。

2 命题的创意和设计

2004年5月在上海复旦大学召开的数学建模工作会议上,全国数学建模竞赛组委会的领导以及众多与会的老师们对竞赛命题发展方向达成共识:提高题目的开放性。这也是我们最初命题时的指导思想。在本题的命题前,我们提出以下立意,并依此设计题目。

1) 贴近实际:通用性和实用性

贴近实际的两层含义中,第一是通用性。而通用性又有两方面的理解,其一是问题是由实际中来的,但问题所适用的不仅限于某个具体问题,而是适用于一类问题,于是问题本身应有比较广泛适用范围。本题的目标虽然是解决奥运会期间的临时商业网点的规模和分布,但是这样的问题适用于更多的场合。例如:春节的游园商业活动,各种赛事及博览会等等;通用性的第二方面理解是对于不同专业的大学生从知识结构上具有通用性。也就是说,各专业的大学生对于题目的内容都易于理解,或适当的参考不太深的一些资料就可以完全理解,这不仅对于不同专业的参赛学生是公平的,而且可以使学生们在紧张的三天赛程中用更多的精力投入在建模工作中,而不是在题目的理解中。这是我们通用性的立意。

贴近实际的第二层含义是实用性。也就是题目从实际中来,又可以应用到实际中去,也就是说题目的设计与表达尽可能地贴近实际。例如,我们从北京规划委那里取来的奥运主场馆的规划图,从中抽取三个主场馆的平面结构图设计题目,周边情况也基本符合真实的情况。同时,我们在题目中提出的三项设计标准:满足供应,分布均衡和商业上盈利,也都是符合国际奥运会要求和符合中国的商业的现实。特别是其中的“分布均衡”是一个很贴近一般人的心理实际的要求。因此,这样设计的题目的好的解答,可能为真实的奥运会提供有用的科学性建议。

2) 提高开放性:浅无边与深无底

在命题的最初,我们就给题目的设计定下这样的开放性要求:解决问题的方法尽可能的广泛,即可以运用各种建模方法和多种数学工具,从初等到高等,从简单到复杂。使不同的数学知识水平的大学生都能在该题的中运用自己的聪明才智,都能得到自己的解答,这就是浅无边的含意。然而即使学生掌握的数学工具十分充分,但对问题的解决也是“天外有天,楼外有楼”的,找到最佳答案是十分困难的,这就是所谓的深无底。

例如:在寻找人流量分布的方法中,可以用初等计算,概率统计方法计算出出行与餐饮过程中人流分布的百分比,也可以用数据挖掘的方法,一下子可以得到比较充分的分布数据。这是一个建立数据模型的过程。而在利用人流规则在计算各商区的人流量的过程中,可以逐个进行累计计算,也可以通过建立网络流图模型去解决。特别是在设计规划问题时,本问题中的三项基本要求:满足供应,分布均衡与商业盈利,是符合实际的条件。特别是其中的分布均衡,要求学生结合实际定义“均衡”,并将均衡性表达成数学形式,这种理解与表达是千差万别的,表现出开放性,这是深无底的一个典型设计。

3) 重在建模:数学模型与数据模型

对于数学建模竞赛来说,建模应当是重点,模型的求解是第二位的。在建模工作中,学生们往往更关注数学建模,例如拟合函数,构建数学规划问题,建立方程与方程组,形成图论问题等等。但是在许多实际问题中,数据模型是经常使用的,也是十分重要的建模工作。因为不是所有问题都可以以简单,典型的方式表达成数学表达式,但是却可以用数据关系把问题的规律清楚地表示出来并付诸应用。因此我们在题目的设计中考虑到把数据模型与数学模型用于同一个题目中。具体说,本题基本上分为两个部分,第一部分是通过对已经给出的问卷调查数据进行分析得到关于出行与餐饮的人员分布的数据模型。这个模型是得到具体的比赛场馆内各商区中人流分布的基础。这个数据模型是用一个数据系统来表达的,模型的完善程度决定了以后

人流分布的确切与详细程度,建立这个模型比较好的方法是用数据挖掘的方法。第二部分是对具体场馆计算商业人流密度与建立网点设计的整数规划问题。其中生成商业人流密度的方法中,比较好的是通过建立网络流模型进行计算。最后的整数规划是一个比较典型的建立规划模型的问题。因此在本题设计中,模型的建立是最重要的工作,而且利用计算模型进行计算也是本题考察的一个重点。

3 阅卷评议

本题的开放性,一方面给大学生发挥自己的创造力提供了比较广阔的空间,同时也对阅卷工作提供了新的课题。

1) 聪明的学生,可喜的成绩

在阅卷中,老师们对该题论文中表现出的富于创造性、严谨的分析论证。聪慧的技巧和善于对自己不熟悉的专业的快速适应能力都留有深刻印象。作为命题人,我对学生能在短短三天内的出色表现叹为观止:今天的大学生前途无量。

本问题的研究过程基本上划分为三个部分:

A. 数据模型的建立

这一部分的目的是通过对三次问卷调查给出的一万条记录的数据进行分析、汇总计算,给出出行与不同类型人流的分布关系,将这些关系数据组成尽可能全面反应相关规律的数据系统,这一目的几乎全部参赛大学生们均很明确。大多数参赛队对三次调查的规律一致性给予充分关注,认为一致性规律才是一般性规律,这是很重要的一步分析。

第二步,在分析与计算的不同的出行(包括进出场馆出行与餐饮出行)与不同类型的人流相关联时,多数采用直观选择可能的相关性进行计算,主要的关系都能计算出,但往往不够完整,其中性别与年龄段对出行方式的考虑不足,由于性别对与出行方式中存在着相关性(例如女性乘出租与私车比例较高),这一条比较容易忽略的规则对计算结果是有影响的。许多论文在这一部分使用了统计相关分析,这是很有用的,结果也比较完整。但是一般的统计方法也需要确定统计计算相关性的对象,这就仍然不能脱离直观的相关属性的选择,也造成不够完善的一个原因。比较少的论文使用了很系统的数据挖掘方法,挖掘出所有二维属性相关值,计算出支持度与置信度,给出完整的数据模型。

B. 人流量分布与购物量分布

数学建模中概念的清楚的定义是很重要的,是否注意到人流量是与购物量是不同的概念是值得首先关注的。在多数论文中都注意到这两个概念的不同,通过购物欲的数据,把人流量转变为购物量。但是其中多数论文没有注意到“人流”本身的定义,因为性别不同和年龄段不同造成出行方式的差别和购物欲的差别。因此,将男女性别的人不加区分地统称为“人”的理解造成了计算上的误差,这一点普遍考虑不足。在计算人流分布(或购物量分布)的方法上,论文表现出最多的创造性。画出路径的网络图,确定最短路径是最普遍使用的方法,许多论文对路口节点的分析是很贴近人们出行与购物习惯的。许多论文利用了矩阵表示商业节点与出行目标之间关系数据,从而使计算变得简便。其中利用到商圈概念时,许多论文学习了商圈理论,使论证有据有力。特别是有些论文十分聪明地构造电路模型或水流模型,用于计算人流分布,这种方法实际上就是网络流模型的一些变形和形象化,也取得很好的效果。一些论文很有特色地引入购物心理学,适当地修正仅用商圈概念的简单模型,应当说是令人心悦诚服的。也有一些论文深刻地研究这一问题,得到一些求人流量的公式,对于更广泛的应用是有意义的。人流分布概率的方法是普遍使用的,应当说是取得好效果的重要方法。值得特别指出的是,多数论文使用了直观的图形与表格进行分析是很重要的方法,这些方法出现在论文中反映了大学生们丰富的创造性和比较广博的知识面。

C. 建立数学规划模型

这最后一部分的目标是明确,绝大多数参赛队都使用数学规划方法,建立数学规划模型。但是,建模的方法很多,思想也各不相同,使得论文丰富多彩。大部分是以商业盈利的要求设计目标函数使其达到最大。另有一部分先计算出每个商区的最大消费额,然后在达到最大消费额条件下求成本最小作为优化目标。在目标函数上选择这两类的论文各有千秋。在建立数学规划模型中,最困难的是如何为满足“均衡”性要求而表达约束条件,这是本题在设计时留下的难点,反映“深无底”的命题特色。在众多论文中主要使用的是限制性约束条件,例如限制在每个商区的MS最多个数与最少个数之差达到极小的约束,这样的规划问题比较简单,也能得到比较符合实际的分布,但是对于商圈数量较大或情况比较复杂的问题,这类约束的想法显得过于简单。

值得注意的是,在一些论文中,一个商区内MS的个数几十甚至上百,对于这样的结果作者们缺乏贴近实际的建模思想,这种不切实际的结果是不应当不加修正就放在论文中的。

D. 论文表达

本题的论文答卷中,最值得一提的是论文的表达,多数论文使用的曲线图、直方图、趋势图、饼图等图形以及用各种数据表表示计算的结果和中间的分析,这不仅十分直观地说明了问题还让判卷老师一目了然地评议,取得了很好的效果。

2) 辛苦的老师,复杂的判卷

开放性赛题必然给老师评卷带来更大的复杂性,这次评卷感受很深。因为开放性题中,浅无边与深无底的要求使得没有“标准答案”,这本身是十分自然的,越贴近实际的问题越可能没有“最佳”结论。评卷工作不仅要求公平、公正,而且还必须有一定的效率。理解一篇论文已经不易,还要论文比较准确的评价,这就对于判卷老师提出更高的要求,不仅要求老师有更广泛的知识而且也必须对题目做更为深入的研究和理解,随着建模题目的开放性的发展,如何能更好地完成判卷工作,或许是一个比较困难的待研究问题。

4 解题思路建议

本题构思立足开放、科学和结合时事。设计原则力求“浅无边,深无底”,也就是说可以用各种方法建模和解题,不同层次的方法都可以使用,得到各种不同解答。特别是,深入分析问题,可以产生科学性很强的构思,很精细的建模和利用相当复杂的数学方法去解决问题。这里给出基本建模和解题的思路。

1) 解题思路

本题基本上可以分成三个部分:从已给问卷调查数据寻找尽可能充分的能够确定人流量的规律并依此计算出每个商区的人流量(人流分布);从人流量分布,提出建立MS的简化假设,形成数学规划问题并解之;评价自己的方法的科学性和结果贴近实际的水平并进行修正。这三个部分应该完整地反映学生的建模构思素质和具体实现能力。

2) 方法建议

A. 首先应该找寻“出行人流规律”与“购物欲规律”。这主要与观众(简化购物主体)的出行目的和购物欲的填表值直接相关。因此,从已经给出的1万多条记录的数据,如何找出尽可能多的出行规律和确定购物欲规律是第一步。方法很多,统计的、数据挖掘的等等,关键是三点原则需要注意:A. 规律是否找出的足够全面,如果用数据挖掘的方法可能比较全面;B. 出行规律中是否都与人流量形成有关?无关的规律应该剔除;C. 三次调查数据规律基本一致的才是普遍规律。

B. 形成人流量分布。有了观众的出行和购物欲规律,又由题目条件,每人每天平均出行两次,而且对出行目标只走最短线,可以计算出20个商区的人流量分布。由于购物欲与性别关系

大,因此假设“标准人”的人流量是较好的方法,例如可以以男性为标准人。人流分布是设计MS网的数据基础,不同方法得到的结果不同,主要是精细程度不同,但是分布比例应当基本一致。但是如果A.的规律得到的结果不够好将直接影响分布比例。计算人流分布的方法很多,建立网络流的计算模型是比较好的方法。

C. 建模假设与数学规划。这是题目开放性的主要体现,假设至少包括两个主要问题:对商店类型的假设,一种类型与两种不同类型,不同类型MS之间在供应能力和利润率的差别。这里主要是确定“满足供应”和“盈利”的数学定义,是建模的关键;第二是对分布的假设,既考虑“分布基本均衡”的要求,例如,不可能因为某区人流量大安排大量的MS,不仅商区面积限制,而且整体不均衡,这种做法是由于没有充分考虑“人是从高密度向低密度流动的”这个基本事实。我们认为扩散模型是表达“均衡性”的比较好的计算模型。例如,沙堆模型就是一种简化的扩散模型,购物人流以沙堆形成的方式由人流量高的商区向临近商区扩散,是一种实现均衡性的模型。当然,必须将这种模型用准确的数学形式表达出来。因此,商业网的模型会是多种多样的,在合理假设之下一般应该形成一个整数规划问题并求解之。

D. 模型的自我评价与修正。这是一个建立模型测试、评价与修改的过程,这项工作的质量,反映学生建模的全面素质,基本原则是建模和解题的科学性,以及是否充分考虑题目中“满足奥运会期间的购物需求、分布基本均衡和商业上赢利”的贴近实际的基本要求。

命题人的具体的题解将另文发表。

5 结束语

有幸参加全国大学生数学建模竞赛的命题和评卷工作,从共事的老师们、专家与领导们身上学习到许多东西,更从大学生们的论文中得到很大启发与鼓舞:我们大学生是十分优秀的。

中国的大学生数模竞赛,2004年参赛队超过六千个,近两万大学生参加数模活动,其规模已经是美国大学生数模竞赛参赛人数的十几倍,堪称最大数模赛事。如何进一步提高和发展这一赛事,使得这一赛事在推动大学生用数学思考世界,用建模解决实际问题的素质和能力的提高是我常常思考的问题。除了大量繁重而复杂的事物性工作以外,命题显然是对这一赛事发展有重大影响的环节。命题的发展趋势,除了贴近时事、提高建模成份以外,大家共识的是提高开放性。然而什么叫开放?这是一个难以定义的概念,有同志提出向美国赛题那样开放,减少命题本身的对操作方法的限制,本文的命题“浅无边深无底”的命题原则也许是属于这样一类开放的尝试。我认为这方面不仅要逐年加强,而且还要更多地探索“美国式”开放以外的开放观念,终究我国的赛事早已在规模上是美国赛事的老大哥,总不能长期以小弟为榜样吧?借此预祝中国的大学生数学建模竞赛命题形成中国式开放模式,把数模竞赛水平推向一个更高的水平上去。

致谢:本文参考了2004年A题的全国一等奖与优秀答卷的论文,同时与叶其孝、姜启源、唐云经过对论文评价的讨论,以及得到孙山泽、王强、谢金星多次相关讨论的帮助在此一并致谢!

Design Temporary Contact Points in Commercial Network About the Main Stadium During the 2008 Beijing Olympic Games

MENG Da-zhi

(Beijing Polytechnic Univ, Beijing 100022)

Abstract: This passage describes simply procedure of putting forward and designing China Undergraduate Mathematical Contest in Modeling in 2004, shows the suggestion of the author's idea of solving problems and original solution, and gives introduction and comment on papers of participants.